BROOKHAVEN
NATIONAL LABORATORY

QUALITY ASSURANCE OF ARM PROGRAM CLIMATE RESEARCH FACILITY DATA. PART II: DATA STREAM INSPECTION AND ASSESSMENT, AND THE ROLES OF VALUE-ADDED PROCESSING AND FIELD CAMPAIGNS IN SPECIFYING DATA QUALITY

Peppler, R. A., Bahrmann, C. P., Doty, K. J., Halter, T. D., Kehoe, K. E., Liljegren, J. C., Long, C. N., Mather, J. H., Moore, S. T., Richardson, S. J., Sisterson, D. L., Sonntag, K. L., Turner, D. D., and Wagener, R.

August 2007

Submitted to
J. Atmos. Oceanic Technol.

Environmental Sciences Department/Atmospheric Sciences Division

Brookhaven National Laboratory

P.O. Box 5000
Upton, NY 11973-5000
www.bnl.gov

Quality assurance of ARM Program Climate Research Facility data. Part II: Data stream inspection and assessment, and the roles of value-added processing and field campaigns in specifying data quality

R. A. Peppler[1], C. P. Bahrmann[2], K. J. Doty[3], T. D. Halter[4], K. E. Kehoe[1], J. C. Liljegren[5], C. N. Long[4], J. H. Mather[4], S. T. Moore[6], S. J. Richardson[2], D. L. Sisterson[5], K. L. Sonntag[1], D. D. Turner[7], and R. Wagener[3]

[1]*NOAA/OAR Cooperative Institute for Mesoscale Meteorological Studies,
The University of Oklahoma, Norman, Oklahoma*

[2]*Pennsylvania State University, State College, Pennsylvania*

[3]*Brookhaven National Laboratory, Upton, New York*

[4]*Pacific Northwest National Laboratory, Richland, Washington*

[5]*Argonne National Laboratory, Argonne, Illinois*

[6]*Mission Research and Technical Services, Santa Barbara, California*

[7]*University of Wisconsin, Madison, Wisconsin*

Submitted to *Journal of Atmospheric and Oceanic Technology*

Address for correspondence: Randy A. Peppler, CIMMS/University of Oklahoma, 2100 National Weather Center, 120 David L. Boren Blvd., Norman, OK 73072-7304.
E-mail: rpeppler@ou.edu

Abstract

This paper, the second in a series of three, documents several key aspects of the Atmospheric Radiation Measurement (ARM) Program Climate Research Facility (ACRF) data quality assurance program. Processes described here include data stream inspection and assessment, and the roles of value-added data processing and the conducting of field campaigns in further specifying data quality and characterizing the basic measurement, respectively. Data collected by field instrumentation must be documented such that users can readily identify whether there are known problems. Data quality inspection and assessment activities have evolved over the 15 years of the ARM Program and have included the development of comprehensive tools and protocols, as provided by instrument mentors, site scientists, and data quality analysts, and their consistent application across ACRF sites. The quality of ACRF instrument-level measurements has been further improved through the processing and analysis of higher-order, value-added data products. These products often provide sophisticated interpretations of measurement-level information that are impossible to make through routine data quality analysis. And, many field campaigns conducted at ACRF sites have included application of observational strategies and instrument comparisons aimed specifically at better measurement characterization; these efforts have led to the improved accuracy and therefore more robust data quality of key ACRF measurements. A comprehensive, end-to-end data quality assurance program is essential for producing a high-quality research data set. The processes developed by the ARM Program offer a possible framework for use by other instrumentation-diverse data-collection networks.

1. Introduction

The performance of ARM Program Climate Research Facility (ACRF) instruments, sites, and data systems is measured in terms of the *availability, usability,* and *accessibility* of the data to a user. Part II in the series describing ACRF data quality assurance processes focuses on the second component, *data usability*. Data must be of sufficient quality for scientific research purposes, and users must be able to readily identify whether there are known problems with the data from quality information embedded both within the data files and contained in ancillary data quality reports. The task of performing data quality checking has evolved over the life of the ARM Program.

Here, we provide a history of ACRF data quality inspection and assessment activities and then describe current tools and practices, including roles and responsibilities, and how data quality information is communicated both to infrastructure staff for problem solving and to the user community for research. We also examine the role of value-added data processing in specifying the quality of the data, and the role of field campaigns in helping better characterize the basic measurement.

2. Evolution of ARM Program data stream checking

a. Early programmatic efforts

Initial data quality assessment was made through the development of self-consistency checks for individual datastreams (Blough 1992) and quality measurement experiments (QMEs; Miller et al. 1994) that compared multiple datastreams. Early self-consistency checking involved not only simple range and rate-of-change tests, but also automated statistical assessment of individual datastreams for internal anomalies to detect outliers and identify instrument failure, creating flags to notify data users and instrument operators. Some statistical assessment was

done using a Bayesian dynamic linear model (DLM). Early applications of DLMs were made

for the detection of moisture on radiometer domes and for the detection of signal attenuation,

side-lobe leakage, birds and other interference on wind profilers.

The QME concept was developed to compare multiple datastreams against a set of

expectations as to the outcome of the comparison, such as the hypothesis of an experiment. The

multiple datastreams that served as inputs to a QME included: direct observations from

instruments; measurements derived from multiple instrument observations and the subsequent

application of algorithms; and model output. The idea behind this concept was that comparisons

involving multiple datastreams should reveal more information about quality than simple single

datastream self-consistency checks can allow. As such, a major function of the QME was to

identify data anomalies and to help analysts identify the root cause of exceptional behavior.   The

defined comparisons were typically made in the near real-time and processing was automated.

The measurements produced by the QME were themselves treated as official ARM data and

were archived. One QME involved comparing vertically integrated water vapor from microwave

radiometers with the output of a microwave radiometer instrument performance model that used

thermodynamic profiles from sondes to drive the model. Another made hourly comparisons

between infrared spectral radiances observed by a Fourier-transform interferometer and the

output of a line-by-line radiative transfer model (Turner et al. 2004b). While within-file flagging

continues today, DLMs and QMEs are no longer processed, though many multi-measurement

and instrument comparisons are still made.

b. Early instrument mentor efforts

Substantial effort was expended in day-to-day quality control early in the program by

instrument mentors, as described in Part I of this series (Peppler 2007a).   In short, mentors

played a crucial role by independently monitoring the data for their assigned instruments using various home-grown analysis and interpretive techniques, reporting their findings on potential problems, and suggesting solutions to site operators. Instrument mentors represented the first line of defense in data quality assessment and problem diagnosis and solution.

c. Early SGP efforts

Southern Great Plains (SGP) site scientists at the University of Oklahoma, after the broadening of the World Wide Web in the mid-1990s, assisted instrument mentors by developing methods to facilitate the graphical display of SGP data and the results of within-file limits checking (Splitt 1996; Peppler and Splitt 1997). The idea behind these diagnostics was to make them available for viewing by instrument mentors and site operators as soon as the data became available (at that time, two days after collection), regardless of the physical location of the viewer. Among the earliest diagnostic aides developed were comparisons between hemispheric broadband solar irradiances and modeled clear sky estimates, and the respective comparisons of shortwave albedo estimates and broadband longwave observations from multiple sites. Interpretive guidance for plots was developed to aid site scientist analysts and an e-mail reporting system was developed for alerting site operators and instrument mentors about problems. Instrument mentors remained fully engaged in near real-time data inspection and assessment activities.

d. Subsequent TWP and NSA efforts

Efforts to display and assess Tropical Western Pacific (TWP) and the North Slope of Alaska (NSA) data quality were made by their respective site scientist teams as data collection commenced later in the 1990s. The TWP locale presented interesting complications that require detailed description here. TWP's first site was installed at Manus Island in October 1996. It

included the core instruments then found at the SGP Central Facility site, including broadband radiometers, a micropulse lidar, and a radiosonde system. But, unlike SGP sites in Oklahoma and Kansas, the available bandwidth connecting Manus to the data management facility (DMF) at the Pacific Northwest National Laboratory in the U.S. was extremely limited. The only satellite communications available were via InmarSat-C for very limited system queries, Inmarsat-B for voice communication, and a GOES data channel. The latter provided the only near real-time data transfer, which amounted to packets of approximately 1 kb each hour. The remainder of the data were recorded to tape and mailed to the U.S., with typical delays of several months between raw data collection and their processing at the DMF.   This delivery delay was the norm until an enhanced VSAT satellite communications system was installed in late 2002, making near-real time delivery possible for most data.

During the period of delayed data delivery from the TWP sites, examination of the data by TWP site scientists then at Pennsylvania State University occurred in two stages. The first stage was to identify potential maintenance issues and was directed at site operations staff. The second stage involved a more detailed review of the data and was directed more toward the science-user community. To address the operations requirement, a compact data status message was constructed that included hourly statistics from most of the instruments along with environmental parameters such as the temperature of the instrument enclosures. These messages were sent via the GOES satellite each hour. Each day, plots of these hourly data were generated and posted on a web site at Penn State. Initially, this process was carried out by the site scientists but was eventually transitioned to site operations. The plots proved useful for identifying gross errors in the data, which fed back to operations team technicians, allowing them to plan repair visits.

Once the full data set was delivered to the DMF, the site scientists produced daily quick looks of these data and performed diagnostic tests, including closure of the solar direct and diffuse components, the net radiative flux, and comparison of the integrated radiosonde water vapor with the vapor derived from the microwave radiometer. Data gaps also were cataloged, which led to uncovering problems with data loggers on several occasions. In the beginning of operation at Manus, data were examined prior to release: once the full data set was received, the data were examined and the site scientist office produced a report describing any issues with the data. This report was submitted to the DMF, though there was not a mechanism at the time to include all of this information to data users. Most of the information was subsequently converted by the DMF and the site scientist office to reports suitable for distribution to users. The procedure of reviewing the data prior to release ended at approximately the time when the second TWP site was installed at Nauru in November 1998. With two sites running, it became impractical to review all of the data prior to release.

During this early phase of operation at the tropical ACRF sites, the site scientist team took a lead role in the examination of the data collected; the initial review of the data was done by a site scientist. When a question regarding a specific instrument arose, the site scientist typically contacted the appropriate instrument mentor and worked with the mentor to solve the problem. Or, if the source of a problem was known or suspected by the site scientist team, TWP operators were contacted to work toward a problem solution. This model of putting the site scientist on the front line of data review had distinct advantages as well as disadvantages. The system had the advantage that the site scientist team had a strong vested interest in the instruments at the site, and looking at multiple instruments provided a holistic view that was useful for problem solving because data from more than one instrument could provide clues

regarding the source of a problem. However, this system was inefficient and time consuming, limiting site scientist time for other important activities like promotion of the data to the science community and planning and implementing field campaigns. With the implementation of a Data Quality Office in 2000, the role of site scientist in reviewing data quality gradually but dramatically changed.

At NSA, still another model was used. Site scientists and site operators jointly subjected data to a systematic program of quality checks (e.g., Delamere et al. 1999). Data streams were visually inspected on a daily basis; from these visual inspections, metadata documenting the overall quality of the data streams were generated. Such inspections facilitated detection of instrument malfunction as the Barrow site was spinning up in 1998. A web-based archive of quick look images was developed to facilitate visual inspection; it was developed and maintained by the site scientists at the Geophysical Institute of the University of Alaska, Fairbanks. These images were updated and made available daily. In addition to visual inspections, the site scientists interpreted within-file limits testing. Instrument mentors had relatively little involvement in quality assurance activities at the NSA after instrument installation and official data release.

NSA site scientists and operators played a crucial role in discussions in 1999 and 2000 on how to better automate data quality checking and how to place this information and other metadata both within NetCDF data files and on the web. Many of these ideas were later adopted by the program after the formation of the Data Quality Office in 2000.

e. Establishment of a Data Quality Office

While instrument mentor and site scientist and site operator efforts were crucial for detecting instrument malfunction and minimizing poor data quality, they tended to be unevenly

developed and applied across the ACRF, sometimes leading to varying treatments of like measurements taken at different locations. A key finding of an ARM Program Infrastructure Review in 1999 stated that "although all groups were well intentioned, there was little coordination of their work. Rather, the effort was unfocused and inconsistent, leading to a general and anecdotal perception both inside and outside ARM that ARM data are less useful than might be expected."[1] A recommendation was made to consolidate and coordinate the program's data quality activities under a central data quality 'czar', which evolved in July 2000 into the Data Quality Office (DQO) at the University of Oklahoma.

The DQO has since coordinated the data quality assurance program, and has particularly helped develop comprehensive tools based on the best of what had been produced by the various contributors to data quality analysis. The DQO has played a key role in ensuring consistent application of data quality protocols across ACRF sites and has broad authority across the ACRF infrastructure to help ensure that the collected data meet the standards required by ARM Program researchers.   The next section documents the results of these efforts. Instrument mentors and site scientists still retain strong, complementary roles in the quality assurance process. Instrument mentors, as the technical authorities for the instruments, still provide in-depth instrument-specific perspectives on data quality and are crucial to resolving problems and identifying long-term data trends. Site scientists, as authorities on their locale and its scientific mission, provide a broad perspective on data quality spanning the full range of site instrumentation. They also oversee their site's problem resolution process and perform targeted research on topics related to site data quality issues. They also directly interact with the scientific community to plan and conduct field campaigns at their sites, which have at times identified previously-unknown data quality issues (see section 6).

[1] ARM Infrastructure Review Committee Report, September 1999, p. 12.

3. Data inspection and assessment

The Data Quality Office is charged with inspecting and assessing approximately 5,000 data fields on a daily to weekly basis. The objective of this activity is to quickly identify and report on data anomalies and provide quick feedback to site operators and instrument mentors to minimize the amount of unacceptable data collected. To perform this task, the DQO has developed, with the technical guidance of instrument mentors and site scientists, automated tools and procedures for inspecting data, all packaged into the Data Quality Health and Status (DQ HandS) system (dq.arm.gov; see for example, Peppler et al. 2005). It facilitates inspection of ACRF datastreams and initiates the problem resolution process, and is built around scripting languages that produce summaries of key parameters. ACRF network configuration allows the DQO to share a file server with the DMF, where raw instrument data from the sites are processed and converted into NetCDF format. This virtual co-location at the data spigot facilitates data quality processing, which allows for quicker problem diagnosis. A DQ HandS prototype was created in the late 1990s as a way to monitor the results of a solar tracker algorithm written by SGP site scientists and was expanded to SGP soil water and temperature system data (Bahrmann and Schneider 1999); from there, it was formalized into a programmatic web-based inspection and assessment tool.

Every hour, the latest available ingested data at the DMF is processed by DQ HandS to create a summary of the bit-packed integer quality control fields (limits tests) within a file. When these fields are not available for an instrument, the DQO works with instrument mentors to identify limits that can be incorporated directly within DQ HandS algorithms. A user interface provides options for selecting the site, datastream, and date range of interest for analysis. After selections are made, a color table showing limits-testing summaries by day is displayed, using a

red/yellow/green color system that bins the data by the percentage passing the limits tests.

Drilling down further provides a more detailed color-coded table by hour for each day, with a mouse-over capability displaying a pop-up window showing more detailed information on flagging results for a particular hour and measurement; statistics include the amount of failure (percentage of collected data) for each automated quality control test violated. Testing that results in flagging now includes internal data stream tests against a set of valid data ranges and comparison tests across data streams against like measurements from different instruments. Color tables are updated hourly as data arrive at the DMF, as are diagnostic plotting aides, which include cross-instrument comparisons. Visual inspection of these plots by DQO analysts helps identify data abnormalities not detectable by automated algorithms. These plots have color-coded backgrounds to indicate local sunlight conditions, helping analysts distinguish between night and day. Fig. 1 shows a representative color table for TWP Manus Island for the first 11 days of May 2007 for a collection of radiometers measuring downwelling radiation, while Fig. 2 shows the corresponding hourly color table for one day (7 May). Fig. 3 displays an example diagnostic plot for 7 May.

A time sequence of daily plots, which can help analysts detect subtle trends in data, is available through a DQ HandS plot browser. Analysts may select a site, an instrument, and a date range, and have the option to view the plot thumbnails for up to 30 days at a time. The thumbnail format is powerful for comparing different instruments that measure like quantities. A user may filter thumbnail results by facility and plot type, and can step forward or backward in time while retaining current filter options. Fig. 4 shows a time sequence of downwelling radiation thumbnails for Manus for 5-8 May 2007.

An interactive, web-based plotting tool called NCVweb also is available in DQ HandS; it allows analysts to easily access and view any NetCDF data at the DMF. Since it works by querying the metadata within each file of interest, there is no need for the data user to be conversant in the NetCDF format. Key features include zooming in on data periods of interest (including less than one day) and plotting multiple data files at one time. Particular data fields of interest can be specified from pull down menus. Plots may consist of one or more independently-plotted fields, multi-dimensional color-coded images such as radar spectra, or slices through a multidimensional array. For closer inspection, data values can be displayed in tabular form or downloaded in ASCII comma-delimited format for easy importation into other applications. Analysts can view file headers to obtain direct access to metadata or can obtain both a summary of data field descriptions and basic field statistics. Fig. 5 shows a zoom in of shortwave direct normal measurement at Manus on 7 May for the hours 0200-1000 UTC, indicating a value that violated a minimum test.

Supplementary data is made available within DQ HandS to assist in data inspection and assessment. Such information includes maintenance field reports from site technicians and data availability statistics from the DMF, plus links to basic information about instrument operation and limitations. The results of data inspections and assessments are issued weekly by the DQO as data quality assessment reports, while data anomalies requiring diagnostic or corrective action require the issuance of problem reports; these reporting mechanisms are described in the next section.

4. Problem reporting, review, and resolution

Data quality reporting mechanisms are based on searchable and accessible databases that allow the various pieces of information produced during the quality assurance process to be

neatly conveyed to problem solvers in a timely manner. The technical design of this database system has evolved with generally available Internet technologies, and as of this writing consists of a Sybase database back end and various CGI scripting languages (PHP, PERL, JavaScript) running on an Apache web-server as the user-accessible front end. Access privileges are controlled by the user's role within the ACRF infrastructure as defined by entries within another database.

The system is divided into four linked subsystems, administered by Brookhaven National Laboratory, with the subsystems based on the nature of the problem: (1) routine weekly assessments of data inspection results are documented through the Data Quality Assessment Report (DQAR) system; (2) routine problems that can be addressed by site operators under the guidance of instrument mentors and site scientists are documented through the Data Quality Problem Report (DQPR) system; (3) significant problems that require engineering effort or that could not be solved in a timely manner through the DQPR process are documented through the Problem Identification Form (PIF)/Corrective Action Report (CAR) system; and (4) the resulting impact on data quality of either problem type is documented by the instrument mentor through the Data Quality Report (DQR) system, which is available to the data user.

The complete history of problems, corrective actions, and reports on data quality is searchable, including criteria for submitter, time period, datastream name, instrument name, designated priority, and assignee. The database allows for the tracking of problem trends and helps identify problematic instrument systems that might require design modifications to make them more reliable and suitable for the continuous operations.

a. Assessment reporting (DQAR subsystem)

To report results of data inspection and assessment, a DQAR is created by DQO analysts

through an interface built into DQ HandS. The site, instrument, and date range selections that

the analyst has been viewing in DQ HandS automatically populate DQAR fields, leaving the

analysts only to type their findings into the form in simple text format. HTML tags can be added

to highlight issues in bold or underline, and hyperlinks can be inserted to provide access to data

plots. The end result is the e-mail issuance of the DQAR to instrument mentors, site scientists,

and site operators. If a data anomaly is discovered during the inspection and assessment process,

the DQAR provides links to problem reports, and pre-populates those reporting forms with

appropriate information (see next subsection). The complete DQAR history is accessible and

searchable through web-based forms by DQAR number, analyst name, date range, site and

instrument.

b. Routine problem reporting (DQPR subsystem)

The DQPR subsystem allows DQO analysts a means to alert appropriate site operators,

site scientists, and instrument mentors about a data anomaly. A powerful feature is its ability to

capture the conversation that documents the progress and status of the diagnostic and corrective

actions proposed and implemented. Once a problem has been identified and a DQPR issued, the

instrument mentor specifies a corrective action and site operators schedule and perform the

proposed action. If additional diagnostic effort is necessary to address the issue, the DQPR

remains open until an alternative solution can be implemented. The DQPR will not be closed

until a corrective action has been deemed successful through the analysis of the quality of

subsequent data. The site scientist, with assistance from the DQO manager, oversees the

progress of problem resolution and each has the responsibility and authority to change problem

status and make work assignments as necessary to resolve an issue in a timely manner. Fig. 6

shows an example DQPR for degraded SGP Vaisala Ceilometer backscatter data caused by a

dirty window. If a problem cannot be resolved through the DQPR process within 30 days, the

DQPR is 'elevated' to a PIF (next subsection), which brings it to the attention of key ACRF

infrastructure personnel.

c. Significant problem reporting (PIF/CAR subsystem)

A PIF may be submitted by anyone involved in the production or use of ACRF data,

including ARM Science Team members and those outside of the ARM Program that may

discover a data anomaly when analyzing data. As previously indicated, site scientists or the

DQO manager can recommend that unsolved DQPR problems be elevated to PIF status, but

sometimes problems are of such magnitude that they require the immediate attention of the

ACRF Problem Review Board (PRB). The PRB consists of senior managers within ACRF

infrastructure representing the Engineering, Operations, Instrument Team, Archive, External

Data, Data System, and DQO groups. It meets once a week via conference call to review new

problems and to track progress on old problems. To facilitate the discussion, an agenda of yet-

to-be reviewed items from the PIF/CAR/DQR database is generated in advance of the call and

the PRB manager prepares disposition recommendations. The PRB reviews each agenda item

and assigns new PIFs to the most appropriate person that can take responsibility for its

resolution. It also assigns a priority to the problem described and specifies an e-mail distribution

of people that need to know about it. The assignee is asked to determine an estimated date of

completion.

The assignee then supervises and monitors problem resolution and submits attachments

to the PIF that serve as progress reports toward correction of the problem. This person also is

responsible for submitting a CAR when the problem has been resolved; the CAR closes out the

problem resolution process with a description of what was done to correct the problem and

triggers the issuance of a DQR (next subsection) to the data user community.

d. Reporting to data users (DQR subsystem)

Any issue having a bearing on the quality of the data, whether reported through a DQPR or a PIF, is ultimately summarized by the instrument mentor in terms of the problem's severity, affected measurements and time period covered. If warranted, suggestions are provided to the data user on how to correct or replace affected data. Such descriptions are captured in a DQR that is distributed by the Data Archive to a data user whenever the affected data are ordered. These reports are worded such that they fully describe the nature of the problem without excessive jargon or unnecessary detail. The theory behind these reports is that they will provide end users with the information they need to make informed judgments on whether and how to use the data. Fig. 7 shows the DQR corresponding to DQPR 1701 in Fig. 6.

5. Role of value-added processing in data quality assurance

Some of the scientific needs of the ARM Program are met through the processing of instrument-level datastreams into value-added products (VAPs). Despite the extensive instrumentation deployed at the ACRF sites, some quantities of interest are either impractical or impossible to measure directly or routinely. VAPs have filled this void and have provided sophisticated interpretations of measurements (e.g., indications of cloud fraction from measurements of solar radiation; estimates of cloud microphysics from radar and lidar data), while at the same time evaluating measurements through the constraints of our understanding of physics (i.e., do the retrieved quantities make sense in the context of the surrounding physical situation?).

VAPs also have shed light on the quality of the data streams used as their input. An example of measurement improvement was accomplished through the Diffuse Correction

(DiffCorr) VAP. Cess et al. (2000) found that some downwelling diffuse shortwave measurements made with shaded Eppley-model precision spectral pyranometers (PSP) under clear-sky conditions fell below the physically-possible limit of diffuse irradiance as produced by a model incorporating both Rayleigh (molecular) scattering and conventional clear-sky atmospheric absorption. Subsequent investigation by Dutton et al. (2001) and Philipona (2002) attributed the problem to infrared (IR) loss from the pyranometer detector causing anomalously low shortwave readings, with Dutton et al. (2001) suggesting a methodology for correcting the measurements using information from co-located longwave pyrgeometer instruments. In implementing the Dutton et al. (2001) methodology as a VAP, Younkin and Long (2004) showed that the PSP IR loss actually exhibits bimodal behavior in the pyrgeometer-pyranometer relationship, dependent on ambient relative humidity conditions, and confirmed the Philipona (2002) findings that the daylight pyranometer IR loss is enhanced compared to that exhibited at night. Thus, the DiffCorr VAP produces an improved measure of downwelling diffuse shortwave irradiance over what the instrument itself is capable of making by correcting for the IR loss inherent in the raw measurements.

Other VAPs have addressed measurement uncertainty. The Best Estimate Flux (BEFlux) VAP (Shi and Long 2002) was designed to produce the best possible measure of surface broadband irradiances for the SGP Central Facility. Instrumentation there includes three separate surface radiation systems located within a few meters of one another. The BEFlux VAP compares the sets of like measurements to test for consistency and then averages the two that agree best to produce a best estimate for that value if that agreement is determined to fall within typical limits established by the historical analysis of known good data. This assessment of historical data then serves as an indication of what range of uncertainty field operations

17

contribute to overall measurement uncertainty. The results of this study are shown in Table 1 (after Shi and Long 2002) as the 95 percent cumulative frequency level of the absolute differences for all data that exhibited the best agreeing pair regardless of which particular instruments were used at any given time (best), the next best agreeing pair (typical), and the worst-agreeing pair (worst). These values, in addition to other factors affecting measurement accuracy (calibration; sensitivity drift between calibrations; contamination of the radiometer domes) can then be used for other single radiometer system sites to set limits on expected performance for quality assessment purposes. In this case, 'best' is the best that can be expected for long-term field measurements, 'typical' is what one might expect on average for ACRF sites, and 'worst' is what one has to consider as the possible operational portion of the total uncertainty of the data.

The range of climatologically-expected values also plays a role in quality assessment. The Quality Control of Radiation (QCRad) VAP (Long and Shi 2006) implements all that the ARM Program has learned in the field about the behavior of surface broadband radiometers and the assessment of their data quality, while at the same time providing a best estimate of their radiation values for data users. For instance, in the development of the DiffCorr VAP (Younkin and Long 2004), considerable effort was expended in the development of methods for testing downwelling longwave measurements, including the pyrgeometer case and dome temperatures and detector fluxes, to prevent the use of questionable pyrgeometer values when correcting for IR loss in the diffuse shortwave measurements. Earlier work for automated quality assessment methodology has been implemented by the international Baseline Surface Radiation Network (BSRN) group (Long and Dutton 2002). The BSRN methodology, which not only sets climatological limits but also makes extensive use of cross comparisons based on known

relationships between variables, has been expanded and improved upon in the QCRad VAP. While the BSRN method uses limits that encompass the entire range of climates from the equator to the poles, the QCRad methodology uses limits based on the particular climatology of a given site, and includes additional tests based on knowledge gained through other VAP development efforts such as the aforementioned BEFlux and DiffCorr, and also the Shortwave Flux Analysis VAP (Long and Ackerman 2000; Long and Gaustad 2004). Thus, VAPs have played a role in not only testing measurements for quality but also in developing methods and expanding knowledge that can be used for improving the testing methodologies themselves.

Testing of measurements using limits and cross-comparisons for quality assurance is fairly standard, but in practice subtle measurement errors that can be difficult to detect using such methods. Additionally, some measurements do not lend themselves to limits testing. In some cases there is no better test for subtle measurement inaccuracies than by using the data in scientific research. One example of this has involved analysis of millimeter cloud radar (MMCR) data; its measured quantity is the reflected radiation from actively broadcast electromagnetic pulses. While the amount of power broadcast and returned can be monitored, there are many other factors involved in this fairly complex instrument that can affect the data quality. The ARSCL (Active Remotely-Sensed Cloud Locations) VAP (Clothiaux et al. 2000; Clothiaux et al. 2001) uses MMCR data as its primary input, and it is within the ARSCL processing that many of the MMCR measurement problems have been revealed. The ARSCL output serves as input into the Baseline Cloud Microphysical Retrievals (MicroBase) VAP (Jensen and Johnson 2006), where again the results are scrutinized in the context of whether the retrievals are consistent with other measurements and are relevant to the physical circumstances in which they are imbedded. Such comparison of parameters in the context of the physical

situation is a powerful tool for determining data quality to a degree not often possible for individual measurements or the retrievals themselves.

As an example of this, MicroBase output is one of the major inputs to the Broadband Heating Rate Profiles (BBHRP) VAP (Mlawer et al. 2002). BBHRP takes the output from several complex VAPs, including ARSCL and MicroBase, and uses it in detailed radiative transfer model calculations. The BBHRP output then is compared with surface and top-of-atmosphere irradiance measurements in a closure experiment framework. It has been through this ongoing model-measurement comparisons that a subtle problem with SGP Central Facility surface direct shortwave measurements was discovered. The comparison revealed a shift in model-measurement agreement statistics for the direct shortwave, which turned out to be caused by human error when two digits of the normal incidence pyrheliometer calibration factor were inadvertently transposed while being entered into a data logger. This human error resulted in a roughly two percent error in the direct shortwave measurements, which turns out to be within the stated uncertainty of the calibrations themselves (Stoffel 2005) and thus was not detectable by standard limits and cross-comparison testing.

In summary, the processing of VAPs has and will continue to serve a significant and valuable purpose in ACRF data quality assurance efforts. This processing also allows the determination of measurement uncertainties and the characterization of instrument behavior, both of which are used to set testing limits in the first place; the development of new procedures to be used in testing; and the improvement in the accuracy of the measured quantity itself by correcting for known errors due to the inherent physical characteristics of the measuring instruments. Through their use in scientific research, VAPs also have provided a powerful tool

for revealing more subtle measurement issues that are difficult or impossible to detect using standard quality control methods.

6. The role of field campaigns in improving data quality

ACRF sites frequently host field campaigns to address specific scientific questions, augment routine data collections, and/or test and validate new instruments. These campaigns are referred to within the ARM Program as intensive observation periods (IOPs). Through 2006, no less than 167 field campaigns have been carried out at the SGP site; 19 campaigns have been held at TWP and 33 have been conducted at NSA. Additionally, eight different campaign activities were held during the inaugural Point Reyes, California, and succeeding Niamey, Niger, ARM Mobile Facility deployments. A specific activity of some campaigns has involved application of observational strategies and instrument comparisons aimed at improving the accuracy and quality of key ACRF measurements. A few key campaigns are described here to show the quality and depth of the work that has been done, which in some cases has had community-wide ramifications on field measurement.

Given the importance of water vapor as a greenhouse gas and its role in the life cycle of clouds and precipitation, the transfer of latent and sensible heat, and atmospheric chemistry, the ARM Program has expended considerable observational effort, particularly at the SGP site, on the measurement of water vapor. Much progress has been made to this end through a series of water vapor IOPs, summarized by Revercomb et al. (2003). These campaigns included three water vapor IOPs held in September 1996, September/October 1997, and September/October 2000, respectively, a lidar IOP held in September/October 1999, and the ARM-First International Satellite Cloud Climatology Project (ISCCP) Regional Experiment (FIRE) Water Vapor Experiment (AFWEX), conducted with NASA in November/December 2000. Additionally, the

Nauru99 campaign held at the TWP ACRF verified microwave radiometer calibration accuracy in the tropical Pacific environment and compared humidity measurements from radiosondes (Westwater et al. 2003).

The 1996 and 1997 water vapor IOPs and the 1999 lidar IOP at SGP provided key information on the quality and accuracy of on-site water vapor instrumentation (Revercomb et al. 2003). Dual-radiosonde launches revealed significant variability across and within calibration batches and showed that differences between any two radiosondes act as an altitude-independent scale factor in the lower troposphere, such that a well-characterized reference can be used to reduce the variability. An approach subsequently was adopted by ARM to scale the radiosonde's moisture profile to agree with the precipitable water vapor observed by a microwave radiometer; this scaling significantly reduced the sonde-to-sonde variability by a factor of two (Turner et al. 2003). The first two water vapor IOPs also were able to verify that 60-m tower-mounted in-situ sensors can serve as an absolute measurement reference, and the site's unique Raman lidar can serve as a stable transfer standard; further IOP results found that the sensitivity of microwave radiometers was excellent over a wide range of integrated water vapor. The third water vapor IOP saw the fielding of further water vapor instrumentation to address leftover issues of absolute calibration.

Also beginning with the first water vapor IOP, verification of site moisture measurements was undertaken through laboratory intercomparison of in situ moisture sensors (including both capacitive chip and chilled mirror sensors) using Oklahoma Mesonet facilities; tests were made both before and after the IOP, making it possible to detect both instrument problems prior to the IOP and instrument failure or drift during the IOP (Richardson et al. 2000). Consequences of

this work were modifications to humidity sensor calibration procedures and the fielding of redundant humidity and temperature sensors to better detect sensor drift and calibration errors.

While the water vapor IOPs were concerned with characterization of water vapor in the lower troposphere, AFWEX was conducted to better characterize upper-tropospheric water vapor measurement (Ferrare et al. 2004). Results from the water vapor IOPs and AFWEX, as described in Soden et al. (2004), showed excellent agreement between satellite and Raman lidar observations of upper tropospheric humidity with systematic differences of about 10 percent; radiosondes, conversely, were found to be systematically drier by 40 percent relative to both satellite and lidar measurements. Existing strategies for correcting the sonde dry bias were found inadequate in the upper troposphere and an alternative method was suggested that considerably improved sonde measurement agreement with lidar observations; it was recommended as a strategy to improve the quality of the global historical record of radiosonde water vapor observations during the satellite era. Further work to characterize the accuracy of Raman lidar water vapor measurements, based on the results of the first two water vapor IOPs, is documented in Turner and Goldsmith (1999), while Ferrare et al. (2006) described evaluation of daytime measurements of water vapor and aerosols made by the Raman lidar during an aerosol IOP conducted in May 2003.

The Nauru99 campaign provided a more subtle understanding of how field campaigns can contribute to data quality by asking how well the data collected accomplished their scientific intent. TWP's equatorial sites make measurements representative of the surrounding oceanic area, which is by far the prevalent surface type. One goal of Nauru99 was to investigate whether the small island of Nauru itself was influencing the measurements made there because of the cloud street phenomenon it produces; this ultimately was confirmed by the campaign. This

campaign result then lead to a year-long Nauru Island Effects Study (NIES; Long 1998), in which a quantification of the island effect on measurements was accomplished, along with the development of a way to detect the effect's ongoing occurrence and estimate its influence on the data (McFarlane et al. 2005). NIES also produced a study that explained the mechanism of cloud street formation (Matthews et al. 2007). Thus, while quantification and detection of the Nauru Island effect is not data quality assessment in the traditional sense of diagnosing instrument performance, it is data quality assessment in the sense of quantifying how well the measurements that are taken address a scientific purpose; here, the island effect is contaminating the goal of making measurements representative of the surrounding oceanic area.

Other important field characterization at the SGP site has been accomplished for measurements of atmospheric radiation. Examples include the second ARM Enhanced Shortwave Experiment (ARESE-II) conducted in February/April 2000 (Michalsky et al. 2002), which focused on broadband shortwave calibration using ground-based and aircraft-mounted radiometers and a standard; a diffuse horizontal shortwave irradiance IOP held in September/October 2001 (Michalsky et al. 2003) that focused attention on a nighttime offset by comparing diffuse irradiance measurements among most commercial pyranometers and some prototypes with the goal of reducing the uncertainty of shortwave diffuse irradiance measurements in lieu of a standard or reference for the measurement; and the first international pyrgeometer and absolute sky-scanning radiometer comparison during September/October 1999 (Philipona et al. 2001), which was conducted to learn more about the reliability and consistency of atmospheric longwave radiation measurements and calculations and to determine their uncertainties, also in lieu of the existence of an absolute standard for the measurement.

Much work also has been done to improve the accuracy of the atmospheric emitted radiance interferometer's (AERI) calibration; this was a focus of the AERI-extended range intercomparison IOP conducted at the NSA site from January 2004 through June 2006 (Turner et al. 2004a). Two instruments were deployed that were identical except that the temperatures of blackbodies used for their calibration were significantly different. This allowed for evaluation of the accuracy of the approach that the ARM Program was using to correct for the non-linear behavior of the AERI's detector. Additionally, during the spectral liquid and ice comparison IOP conducted at the SGP site in October 2003, a second AERI was deployed running in a prototype "rapid-sampling" mode (i.e., sky samples every 20-30 s instead of every 7-8 min); this side-by-side comparison with the routine operational AERI running a different temporal sampling strategy allowed scientists to assess the adequacy of a newly-developed noise filter (Turner et al. 2006). On the strength of this experiment, ARM was convinced to adopt the new rapid-sampling mode in all of its AERIs.

7. Summary

Part II describing the ARM Program Climate Research Facility data quality assurance program has focused on data quality inspection and assessment and the important roles of value-added data processing and the conducting of field campaigns in specifying data quality and measurement accuracy.

Data collected by field instrumentation must be documented such that users can readily identify whether there are known problems. Data quality inspection and assessment activities have evolved over the 15-year life of the ARM Program, culminating in the formation of a Data Quality Office in July 2000. The DQO has since coordinated the data quality program, including the development of comprehensive inspection, assessment, and reporting tools, and data

checking protocols, based on the best of what had been produced by the various players in data quality up to that time. The DQO has played a key role in ensuring consistent application of data quality protocols across ACRF sites. Instrument mentors, as the technical authorities for the instruments, provide in-depth instrument-specific perspectives on data quality and are crucial to resolving problems and identifying long-term data trends. Site scientists, as authorities on the locale and the site's scientific mission, provide a broad perspective on data quality spanning the full range of instrumentation and oversee problem resolution. They also may perform targeted research on topics related to site data quality issues, and have the vital role of interacting with the scientific community to plan and conduct field campaigns at their sites, which have at times have identified previously-unknown data quality issues.

The quality of ACRF instrument-level measurements has been further improved through the processing and analysis of higher-order, value-added data products. These products often provide sophisticated interpretations of measurement-level information that are impossible to make through routine data quality analysis. And, many field campaigns have involved application of observational strategies and instrument comparisons aimed specifically at better measurement characterization; these have led to the improved accuracy and therefore more robust data quality of key ACRF measurements.

Part III (Peppler et al. 2007b) on ACRF data quality assurance will describe data archival, display and distribution, data stream reprocessing, engineering and operations management processes and procedures, and future directions in data quality assurance, while Part I (Peppler et al. 2007a) described instrument deployment, instrument and facility maintenance, and data collection and processing infrastructure, along with a brief overview of past work on data quality assurance and ARM Program history. The processes we have developed represent a possible

framework for use by other instrumentation-diverse networks; we offer them to the community in this three-part series to highlight the myriad aspects that go into producing a robust, high-quality data set suitable for scientific research.

References

Bahrmann, C. P., and J. M. Schneider, 1999: Near real-time assessment of SWATS data quality, resulting in an overall improvement in present-day SWATS data quality. Proc., *Ninth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, available at http://www.arm.gov/publications/proceedings/conf09/extended_abs/bahrmann_cp.pdf.

Blough, D. K., 1992: Real-time statistical quality control and ARM. Transactions, *46th Annual ASQC Quality Congress*, Nashville, TN, Amer. Soc. Qual. Ctrl., 484-490.

Cess, R. D., T. Quian, and M. Sun, 2000: Consistency tests applied to the measurement of total, direct and diffuse shortwave radiation at the surface. *J. Geophys. Res.-Atmospheres*, **105**, 24881-24887.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.*, **39**, 645-665.

Clothiaux, E. E., M. A. Miller, R. C. Perez, D. D. Turner, K. P. Moran, B. E. Martner, T. P. Ackerman, G. G. Mace, R. T. Marchand, K. B. Widener, D. J. Rodriguez, T. Uttal, J. H. Mather, C. J. Flynn, K. L. Gaustad, and B. Ermold, 2001: The ARM Millimeter Wave Cloud Radars (MMCRs) and the Active Remote Sensing of Clouds (ARSCL) Value Added Product (VAP). DOE Tech. Memo. *ARM VAP-002.1*, U.S. Department of Energy, Washington, D.C., available at http://www.arm.gov/publications/tech_reports/armvap-002-1.pdf.

Delamere, J. S., H. A. Eide, H. Lindquist, R. Storvold, P. Utley, K. Stamnes, C. R. Turney, K. B. Widener, and B. Zak, 1999: The first year of operation of the North Slope of Alaska/Adjacent Arctic Ocean ARM site: An overview of instrumentation, data streams, and data quality assurance procedures. Proc., *Ninth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, available at http://www.arm.gov/publications/proceedings/conf09/extended_abs/delamere1_js.pdf.

Dutton, E. G., J. J. Michalsky, T. Stoffel, B. W. Forgan, J. Hickey, D. W. Nelson, T. L. Alberta, and I. Reda, 2001: Measurement of broadband diffuse solar irradiance using current commercial instrumentation with a correction for thermal offset errors. *J. Atmos. Oceanic Technol.*, **18**, 297–314.

Ferrare, R. A., E. V. Browell, S. Ismail, S. Kooi, L. H. Brasseur, V. G. Brackett, M. B. Clayton, J. D. W. Barrick, G. S. Diskin, J. E. M. Goldsmith, B. M. Lesht, J. R. Podolske, G. W. Sachse, F. J. Schmidlin, D. D. Turner, D. N. Whiteman, D. Tobin, L. Miloshevich, H. E. Revercomb, B. B. Demoz, and P. DiGirolamo, 2004: Characterization of upper tropospheric water vapor measurements during AFWEX using LASE. *J. Atmos. Oceanic Technol.*, **21**, 1790-1808.

Ferrare, R., D. Turner, M. Clayton, B. Schmid, J. Redemann, D. Covert, R. Elleman, J. Ogren, E. Andrews, J. E. M. Goldsmith, and H. Jonsson, 2006: Evaluation of daytime measurements of aerosols and water vapor made by an operational Raman lidar over the Southern Great Plains. *J. Geophys. Res.-Atmospheres*, **111** (D5), Art. No. D05S08.

Jensen, M., and K. Johnson, 2006: Continuous Profiles of Cloud Microphysical Properties for the Fixed Atmospheric Radiation Measurement Sites. Atmospheric Radiation Measurement Program Technical Report *DOE/SC-ARM/P-0609*, 7 pp., available at http://www.arm.gov/publications/programdocs/doe-sc-arm-p-06-009.pdf.

Long, C. N., 1998: Nauru Island Effect Study (NIES) IOP Science Plan. ARM Technical Document *DOE/SC-ARM-0505*, 17 pp., available at http://www.arm.gov/publications/programdocs/doe-sc-arm-0505.pdf.

Long, C. N., and T. P. Ackerman, 2000: Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. *J. Geophys. Res.-Atmospheres*, **105** (D12), 15609-15626.

Long, C. N., and E. G. Dutton, 2002: BSRN Global Network recommended QC Tests, V2.0. BSRN Technical Report, available at http://ezksun3.ethz.ch/bsrn/admin/dokus/qualitycheck.pdf.

Long, C. N., and K. L. Gaustad, 2004: The Shortwave (SW) Clear-Sky Detection and Fitting Algorithm: Algorithm Operational Details and Explanations. Atmospheric Radiation Measurement Program Technical Report *ARM TR-004.1*, 26 pp., available at http://www.arm.gov/publications/tech_reports/arm-tr-004-1.pdf.

Long, C. N., and Y. Shi, 2006: The QCRad Value Added Product: Surface Radiation Measurement Quality Control Testing, Including Climatologically Configurable Limits. Atmospheric Radiation Measurement Program Technical Report *ARM TR-074*, 69 pp., available at http://www.arm.gov/publications/tech_reports/arm-tr-074.pdf.

Matthews, S., J. M. Hacker, J. Cole, J. Hare, C. N. Long, and R. M. Reynolds, 2007: Modification of the atmospheric boundary layer by a small island: Observations from Nauru. *Mon. Wea. Rev.*, **135**, 891-905.

McFarlane, S. A., C. N. Long, and D. M. Flynn, 2005: Impact of island-induced clouds on surface measurements: Analysis of the ARM Nauru Island Effect Study data. *J. Appl. Meteor.*, **44**, 1045-1065.

Mlawer, E. J., J. S. Delamere, S. A. Clough, M. A. Miller, K. L. Johnson, T. R. Shippert, C. N. Long, R. G. Ellingson, M. H. Zhang, R. A. Ferrare, and R. T. Cederwall, 2002: The Broadband Heating Rate Profile (BBHRP) VAP. Proc., *12th ARM Science Team Meeting*, St. Petersburg, FL, available at http://www.arm.gov/publications/proceedings/conf12/extended_abs/mlawer-ej.pdf.

Michalsky, J. J., R. Dolce, E. G. Dutton, M. Haeffelin, G. Major, J. A. Schlemmer, D. W. Slater, J. R. Hickey, W. Q. Jeffries, A. Los, D. Mathias, L. J. B. McArthur, R. Philipona, I. Reda, and T. Stoffel, 2003: Results from the first ARM diffuse horizontal shortwave irradiance comparison. *J. Geophys. Res.-Atmospheres*, **108** (D3), Art. No. 4108.

Michalsky, J., P. Kiedron, J. Berndt, T. Stoffel, D. Myers, I. Reda, J. Treadwell, A. Andreas, S. Asano, A. Uchiyama, A. Yamazaki, M. Haeffelin, T. Tooman, R. McCoy, A. Bucholtz, B. C. Bush, S. K. Pope, A. S. Leitner, and F. P. J. Valero, 2002: Broadband shortwave calibration results from the Atmospheric Radiation Measurement Enhanced Shortwave Experiment II. *J. Geophys. Res.-Atmospheres*, **107** (D16), Art. No. 4287.

Miller, N. E., J. C. Liljegren, T. R. Shippert, S. A. Clough, and P. D. Brown, 1994: Quality measurement experiments within the Atmospheric Radiation Measurement Program. Preprints, *Fifth Symp. on Global Change Studies*, Nashville, TN, Amer. Meteor. Soc., 35-39.

Peppler, R. A., and M. E. Splitt, 1997: SGP Site Scientist Team data quality assessment activities. Proc., *Seventh Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 403-406.

Peppler, R. A., K. E. Kehoe, K. L. Sonntag, S. T. Moore, and K. J. Doty, 2005: Improvements to and status of ARM's Data Quality Health and Status System. Preprints, *15th Conf. on Applied Climatology*, Savannah, GA, Amer. Meteor. Soc., CD-ROM, J3.13.

Peppler, R. A., C. P. Bahrmann, R. C. Eagan, M. D. Ivey, N. N. Keck, J. C. Liljegren, C. N. Long, M. C. Macduff, J. H. Mather, K. L. Nitschke, B. W. Orr, S. J. Richardson, D. L. Sisterson, and J. W. Voyles, 2007a: Quality assurance of ARM Program Climate Research Facility data. Part I: Instrument and facility deployment and maintenance, data collection and processing infrastructure, and engineering and operations management. *J. Atmos. Oceanic Technol.*, **24**, submitted.

Peppler, R. A., S. W. Christensen, K. E. Kehoe, C. N. Long, R. A. McCord, S. T. Moore, R. C. Perez, B. D. Perkins, D. L. Sisterson, K. L. Sonntag, and J. W. Voyles, 2007b: Quality assurance of ARM Program Climate Research Facility data. Part III: Data archival, display and distribution, data stream reprocessing, and future directions in data quality assurance. *J. Atmos. Oceanic Technol.*, **24**, submitted.

Philipona, R., 2002: Underestimation of solar global and diffuse radiation measured at Earth's surface. *J. Geophys. Res.-Atmospheres*, **107** (D22), Art. No. 4654.

Philipona, R., E. G. Dutton, T. Stoffel, J. Michalsky, I. Reda, A. Stifter, P. Wendling, N. Wood, S. A. Clough, E. J. Mlawer, G. Anderson, H. E. Revercomb, and T. R. Shippert, 2001: Atmospheric longwave irradiance uncertainty: Pyrgeometers compared to an absolute sky-scanning radiometer, atmospheric emitted radiance interferometer, and radiative transfer model calculations. *J. Geophys. Res.-Atmospheres*, **106** (D22), 28129-28141.

Revercomb, H. E., D. D. Turner, D. C. Tobin, R. O. Knuteson, W. F. Feltz, J. Barnard, J. Bösenberg, S. Clough, D. Cook, R. Ferrare, J. Goldsmith, S. Gutman, R. Halthore, B. Lesht, J. Liljegren, H. Linné, J. Michalsky, V. Morris, W. Porch, S. Richardson, B. Schmid, M. Splitt, T. Van Hove, E. Westwater, and D. Whiteman, 2003: The ARM Program's water vapor intensive observation periods. *Bull. Amer. Meteor. Soc.*, **84**, 217-236.

Richardson, S. J., M. E. Splitt, and B. M. Lesht, 2000: Enhancement of ARM surface meteorological observations during the fall 1996 water vapor intensive observation period. *J. Atmos. Oceanic Technol.*, **17**, 312-322.

Shi, Y., and C. N. Long, 2002: Best Estimate Radiation Flux Value Added Product: Algorithm Operational Details and Explanations. Atmospheric Radiation Measurement Program Technical Report *ARM TR-008*, 55 pp., available at http://www.arm.gov/publications/tech_reports/arm-tr-008.pdf.

Soden, B. J., D. D. Turner, B. M. Lesht, and L. M. Miloshevich, 2004: An analysis of satellite, radiosonde, and lidar observations of upper tropospheric water vapor from the Atmospheric Radiation Measurement Program. *J. Geophys. Res. – Atmospheres*, **109** (D4), Art. No. D04105.

Splitt, M. E., 1996: Data quality display modules – assessment of instrument performance at the Southern Great Plains Cloud and Radiation Testbed site. Proc., *Sixth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 301-303.

Stoffel, T., 2005: Solar Infrared Radiation Station (SIRS) Handbook. Atmospheric Radiation Measurement Program Technical Report *ARM TR-025*, 27 pp., available at http://www.arm.gov/publications/tech_reports/handbooks/sirs_handbook.doc.

Turner, D. D., and J. E. M. Goldsmith, 1999: Twenty-four-hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement Program's 1996 and 1997 water vapor intensive observation periods. *J. Atmos. Oceanic Technol.*, **16**, 1062-1076.

Turner, D. D., R. O. Knuteson, H. E. Revercomb, C. Lo, and R. G. Dedecker, 2006: Noise reduction of Atmospheric Emitted Radiance Interferometer (AERI) observations using principal component analysis. *J. Atmos. Oceanic Technol.*, **23**, 1223-1238.

Turner, D. D., B. M. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin, 2003: Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience. *J. Atmos. Oceanic Technol.*, **20**, 117-132.

Turner, D. D., H. E. Revercomb, R. O. Knuteson, R. G. Dedecker, and W. F. Feltz, 2004a: An evaluation of the nonlinearity correction applied to Atmospheric Emitted Radiance Interferometer (AERI) data collected by the Atmospheric Radiation Measurement

Program. Atmospheric Radiation Measurement Program Technical Report *ARM TR-013*, 7 pp., available at http://www.arm.gov/publications/tech_reports/arm-tr-013.pdf.

Turner, D. D., D. C. Tobin, S. A. Clough, P. D. Brown, R. G. Ellingson, E. J. Mlawer, R. O. Knuteson, H. E. Revercomb, T. R. Shippert, and W. L. Smith, 2004b: The QME AERI LBLRTM: A closure experiment for downwelling high spectral resolution infrared radiance. *J. Atmos. Sci.*, **61**, 2657-2675.

Westwater, E. R., B. B. Stankov, D. Cimini, Y. Han, J. A. Shaw, B. M. Lesht, and C. N. Long, 2003: Radiosonde humidity soundings and microwave radiometers during Nauru99. *J. Atmos. Oceanic Technol.*, **20**, 953-971.

Younkin, K., and C. N. Long, 2004: Improved Correction of IR Loss in Diffuse Shortwave Measurements: An ARM Value Added Product. Atmospheric Radiation Measurement Program Technical Report *ARM TR-009*, 50 pp., available at http://www.arm.gov/publications/tech_reports/arm-tr-009.pdf.

Tables

Table 1: Level at which 95 percent cumulative frequency of agreement was reached between radiometer measurements (after Shi and Long 2002).

Measurement:	Best	Typical	Worst
Diffuse SW	4.0 ± 1.4	9.0 ± 3.1	12.0 ± 3.8
Direct Normal SW	6.3 ± 3.3	13.6 ± 6.3	15.0 ± 6.7
Downwelling LW	3.1 ± 0.4	5.1 ± 1.2	7.1 ± 1.4
Upwelling SW	11.1 ± 2.8		
Upwelling LW	9.6 ± 3.0		

Figure Captions

Fig. 1. Daily color table of automated quality control check results for TWP Manus Island site for 1-11 May 2007 for a collection of radiometers that measure downwelling shortwave radiation.

Fig. 2. Hourly color table of automated quality control check results for one day (7 May 2007) at the TWP Manus Island site corresponding to the example in Fig 1. Blue pop-up window of flagging statistics is obtained by mousing over the yellow shaded box for 0500 UTC for the shortwave direct normal incidence measurement; here, 7% of the observations failed a minimum test.

Fig. 3. Diagnostic plot of some downwelling shortwave radiation variables on 7 May 2007 at the TWP Manus Island site corresponding to the example in Figs 1-2. 'sdn' is the shortwave direct normal component; 'dsdh' is downwelling shortwave diffuse hemispheric component, and 'dsh' is the downwelling shortwave hemispheric measurement. 'Derived' hemispheric is calculated by adding the diffuse radiation component (measured by a shaded pyranometer) to the direct normal radiation component (measured by a pyrheliometer) multiplied by the cosine of the zenith angle; it is then compared as a quality control check to the actual hemispheric measurement provided by an unshaded pyranometer (denoted as 'dsh'). The yellow vertical dashed line denotes local solar noon.

Fig. 4. Plot browser time sequence of downwelling radiation plots for 5-8 May 2007 corresponding to the example in Figs. 1-3 for TWP Manus Island.

Fig. 5. NCVweb zoom-in on the hours 0200-1000 UTC of the shortwave direct normal measurement on 7 May 2007 corresponding to the example in Figs. 1-4, showing values (red asterisk) denoted in Fig. 2 that violated a minimum test.

Fig. 6. Data Quality Problem Report (DQPR) 1701 identifying degraded Vaisala Ceilometer backscatter data at SGP Boundary Facility 6 in Purcell, Oklahoma, that ultimately were found to be caused by a dirty window. Comments reveal the series of steps taken to rectify and report on the problem.

Fig. 7. Data Quality Report (DQR) D070219.1 issued to data users to alert them to the issue described in DQPR 1701 (see Fig. 6). This DQR is attached to any data order containing this particular measurement and time period.

Figures

Fig. 1. Daily color table of automated quality control check results for TWP Manus Island site for 1-11 May 2007 for a collection of radiometers that measure downwelling shortwave radiation.
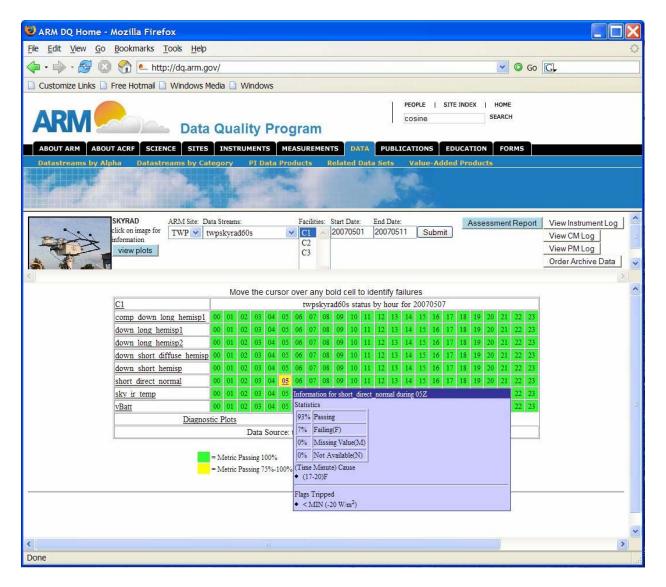
Fig. 2. Hourly color table of automated quality control check results for one day (7 May 2007) at the TWP Manus Island site corresponding to the example in Fig. 1. Blue pop-up window of flagging statistics is obtained by mousing over the yellow shaded box for 0500 UTC for the shortwave direct normal incidence measurement; here, 7% of the observations failed a minimum test.

Fig. 3. Ddiagnostic plot of some downwelling shortwave radiation variables on 7 May 2007 at the TWP Manus Island site corresponding to the example in Figs. 1-2. 'sdn' is the shortwave direct normal component; 'dsdh' is downwelling shortwave diffuse hemispheric component, and 'dsh' is the downwelling shortwave hemispheric measurement. 'Derived' hemispheric is calculated by adding the diffuse radiation component (measured by a shaded pyranometer) to the direct normal radiation component (measured by a pyrheliometer) multiplied by the cosine of the zenith angle; it is then compared as a quality control check to the actual hemispheric measurement provided by an unshaded pyranometer (denoted as 'dsh'). The yellow vertical dashed line denotes local solar noon.
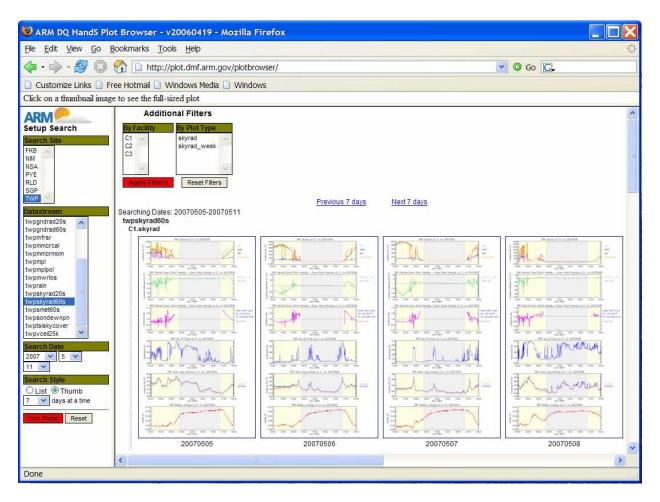
Fig. 4. Plot browser time sequence of downwelling radiation plots for 5-8 May 2007
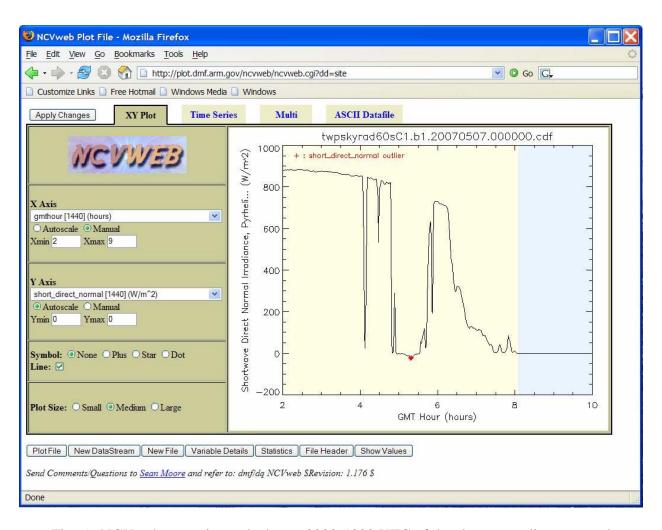corresponding to the example in Figs. 1-3 for TWP Manus Island.

Fig. 5. NCVweb zoom-in on the hours 0200-1000 UTC of the shortwave direct normal measurement on 7 May 2007 corresponding to the example in Figs. 1-4, showing values (red asterisk) denoted in Fig. 2 that violated a minimum test.
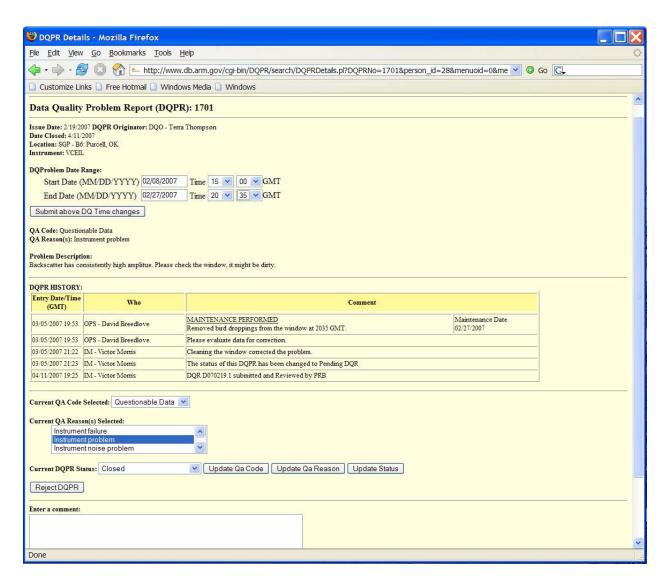
Fig. 6. Data Quality Problem Report (DQPR) 1701 identifying degraded Vaisala Ceilometer backscatter data at SGP Boundary Facility 6 in Purcell, Oklahoma, that ultimately were found to be caused by a dirty window. Comments reveal the series of steps taken to rectify and report on the problem.
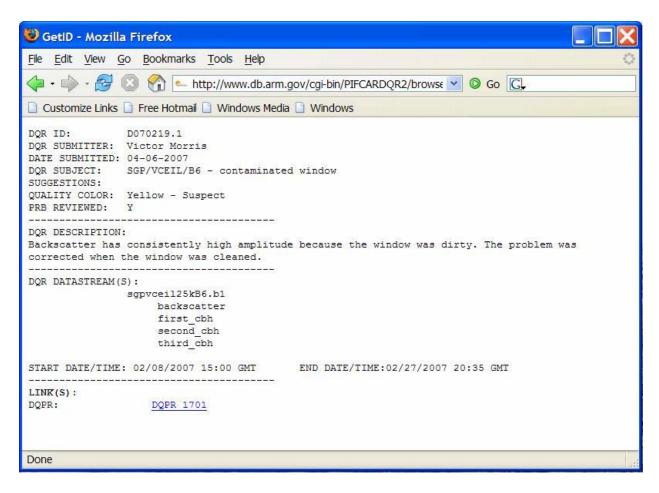
Fig. 7. Data Quality Report (DQR) D070219.1 issued to data users to alert them to the issue described in DQPR 1701 (see Fig. 6). This DQR is attached to any data order containing this particular measurement and time period.